Accelerating Protein Docking Calculations using the ATTRACT Coarse-Grained Force Field and 3D Rotation Maps

Maria Elisa Ruiz-Echartea[1, 3], Isaure Chauvot de Beauchêne[2], David W Ritchie [3]

1. Université de Lorraine, Campus Scientifique, Vandoeuvre-lès-Nancy, 54500 France
2. CNRS, LORIA, Campus Scientifique, Vandoeuvre-lès-Nancy, 54500 France
3. INRIA Nancy – Grand Est, 615 Rue du Jardin Botanique, Villers-lès-Nancy, 54600 France

Introduction

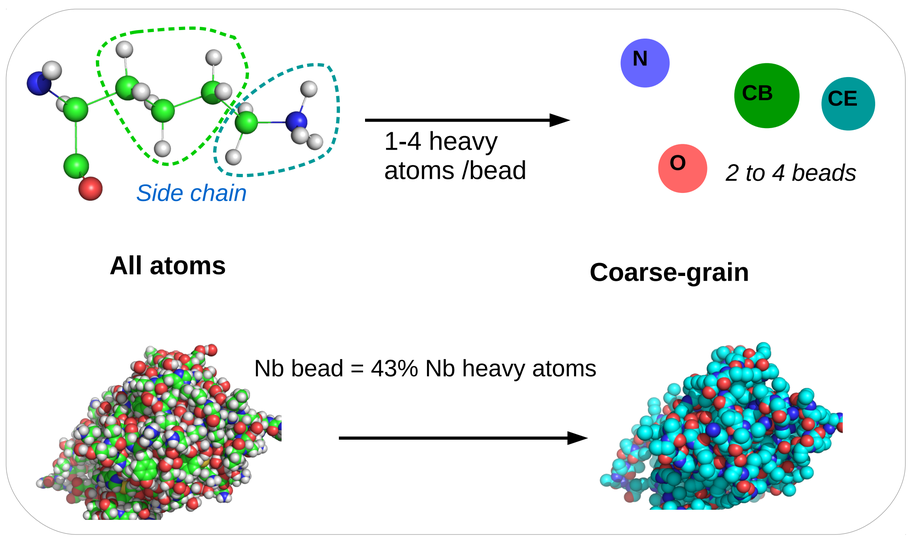
Protein-protein docking algorithms aim to predict how two proteins interact with each other to form a complex. Docking algorithms need to fulfil two main tasks: (1) sampling the possible relative positions of the two proteins and (2) computing the interaction energy at each position. Although the docking problem has been studied for over 25 years, doing this accurately and thoroughly remains a computationally expensive task. We are therefore developing a new algorithm to face the protein docking problem in a novel and efficient way.

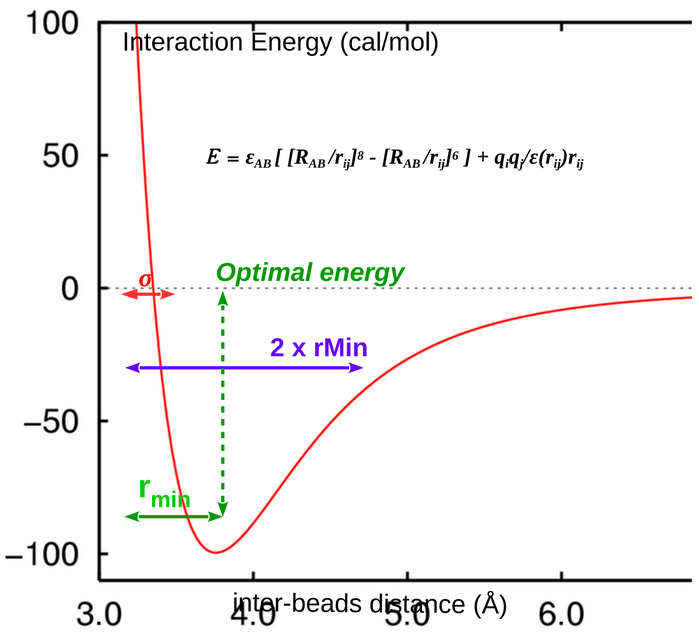
We are interested in performing an exhaustive search of the rigid-body docking space, where the positioning of the two proteins is not driven by any prior knowledge (e.g. from homology), while still using an accurate force field interaction energy model. In order to reduce the O(N^2) cost of atomistic force-field models, we use a coarse-grained (CG) bead model taken from ATTRACT [1]. However, a naive energy calculation for every trial orientation would still cost O(N^2) energy evaluations in the number of beads. We are therefore developing a method to detect the locations of all possible clashes before performing any energy calculations, thus allowing us to avoid calculating energies for many millions of useless trial orientations.

Based on a preliminary study of protein-protein interfaces in complexes from the Protein Docking Benchmark [2], we found that a large number of interfaces contain at least one pair of CG beads at almost their optimal distance. Therefore, our idea is to perform a series of restricted docking searches in which one surface bead from each docking partner is placed in contact at the coordinate origin. This leaves a 3D rotational search, in which ligand may rotate around a fixed receptor.

Of course, a full docking search requires all possible pairs of surface beads to be placed together. However, within each rotational sub-problem, we can exploit the fact that rotations do not change any distances from the origin. Thus, we can pre-calculate a "3D rotational map" of all of the rotations that will cause the beads to clash. We can then restrict the remaining rotational search and energy calculation to a small region near the forbidden rotations in the clash map.

ATTRACT Coarse-Grained Energy Model



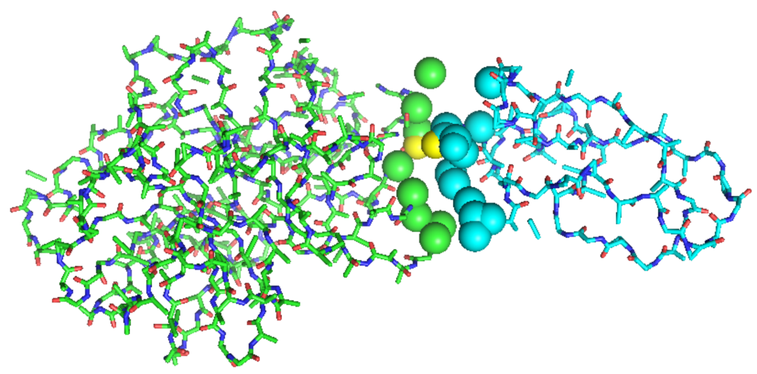


The Protein Docking Benchmark v5.0

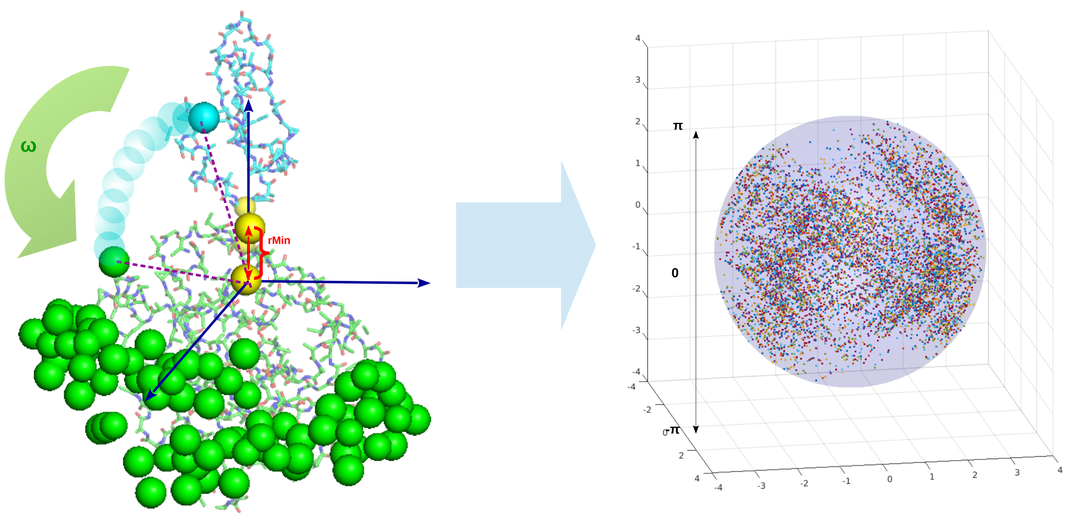
We looked at all of the protein-protein complexes in the Protein Docking Benchmark (v5) in order to check how often it is possible to find at least one pair of coarse-grained beads located at their optimal distance (shown in yellow on the right), namely, pairs of beads which have the lowest possible ATTRACT interaction energy. To do this, we computed the difference between the best bead-bead distance in the complex and compared it to the optimal distance for the corresponding bead types (deltaD). The results are shown in the table below.

Complex Type	deltaD	% of complexes
Bound	< 0.1 Å	93
	< 0.23 Å	100
Unbound fitted	< 0.1 Å	86
	< 0.2 Å	98

Each complex has between 3 and 80 near-optimal bead pairs, and as this table shows 98% of the unbound complexes has at least one pair of beads within 0.2 Å of the optimal separation.



Detecting Steric Clashes – 3D Clash Maps



For the translational part of a docking search, we consider all possible pairs of receptor and ligand surface beads in turn, and for each pair we translate the ligand and receptor so that the selected beads (here shown in yellow) are in perfect contact at the origin with their centres of mass are aligned on the z-axis. This absorbs 3 translational degrees of freedom, and leaves a pure 3D rotational search problem about the origin for each selected pair of beads, as illustrated on the left.

Before calculating ATTRACT bead energies, we use the bead coordinates to calculate a list of rotations that will cause steric clashes. These are rotations that should be avoided during the docking search. A 3D representation of this "clash map" is shown on the right. It can be seen that only a small region of the search space is not forbidden by steric clashes.

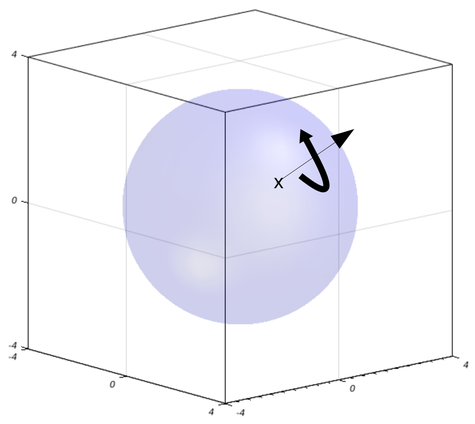
3D Rotation Sampling and ATTRACT Energy Calculation

Each 3D docking search is performed by recursively sub-dividing a 3D quaternion "π-ball" representation of the rotational space. Rotational samples generated in this way may be rapidly compared to rotations in the 3D clash map.

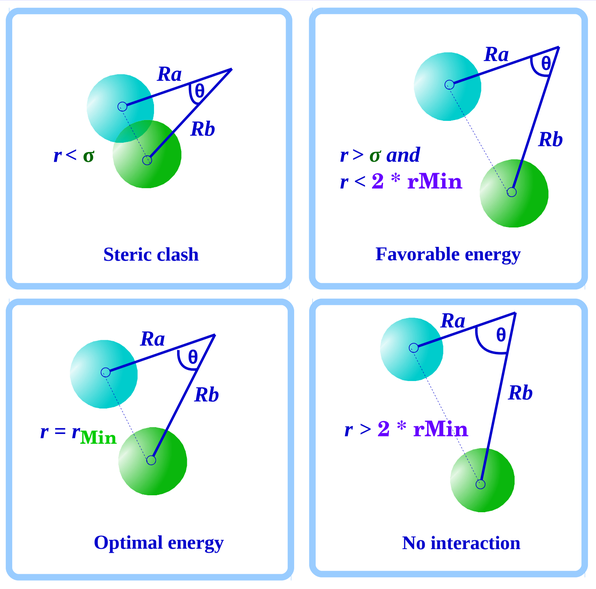
Using each bead's distance from the origin (here, R_a and R_b), the cosine rule

$$r^2 = R_a^2 + R_b^2 - 2R_a R_b \cos\theta$$

may be used to calculate pair-wise bead distances very efficiently from a quaternion rotation (details not shown). Thus, we can detect steric clashes and calculate interaction energies without ever needing to physically move any beads.



Each point in the π-ball defines a rotation axis and angle. A rotational search is performed by recursively sub-dividing the π-ball.



Conclusions

We have found that many protein-protein complexes have at least one pair of beads in near-optimal contact. Thus, using pairs of surface beads to initialise a docking search is a valid heuristic for treating 3D translations.

The algorithm described here will allow us to perform an exhaustive recursive docking search of 3D rotation space. Thus, in principle, our approach will never miss any local energy minima, and is guaranteed to find the global minimum of the ATTRACT energy function for each pair of starting beads.

Our algorithm is still under development. Nonetheless, we believe it could be faster than the conventional gradient minimisation that is currently used in ATTRACT. It also promises to be more accurate if we can train a new scoring function to optimise the local minima among the sampled poses.

Eventually, we will test our approach on the Protein Docking Benchmark.

References

1. Zacharias, M.
Protein Science, 12, 1271-1282 (2003).

2. Hwang, H., Vreven T., Janin, J., Weng Z.
Proteins, 78, 3111-3114 (2010).

Acknowledgements

Maria Elisa Ruiz Echartea is supported by a doctoral bursary (CORDI-S) from Inria Nancy Grand Est.